A TEXTURE-BASED METHODOLOGY FOR IDENTIFYING TISSUE TYPE IN MAGNETIC RESONANCE IMAGES

Michael Barnathan[1], Jingjing Zhang[1], Erickson Miranda[1], Vasileios Megalooikonomou[1], Scott Faro[2], Harvey Hensley[3], Luis Del Valle[4], Kamel Khalili[4], Jennifer Gordon[4], Feroze B. Mohamed[2]

[1]Data Engineering Laboratory (DEnLab), Department of Computer and Information Sciences, Temple University, 1805 N. Broad St. Philadelphia, PA 19122

[2]Department of Radiology, Temple University School of Medicine, 3401 N. Broad St. Philadelphia, PA 19140

[3]Fox Chase Cancer Center, 333 Cottman Ave. Philadelphia, PA 19111

[4]Department of Neurovirology and Cancer Biology, Temple University, 1900 N. 12th St. Philadelphia, PA 19122

ABSTRACT

We propose a methodology for discriminating between various types of normal and diseased brain tissue in medical images that utilizes Vector Quantization (VQ), an image compression technique, to extract discriminative texture features. Rather than focusing on images of the entire brain, we direct our attention to extracting local descriptors for individual regions of interest (ROIs) as determined by domain experts. After determining regions of interest, we generate a "locally optimal" codebook representing texture features of each region using the Generalized Lloyd Algorithm. We then utilize the codeword usage frequency of each codeword in the codebook as a discriminative feature vector for the region it represents. Finally, we compare k-nearest neighbor, neural network, support vector machine, and decision tree-based classification approaches using the Histogram Model (HM) distance metric. Combined T1 and T2 classification accuracies in mice averaged 89% under certain experimental settings, indicating that our approach may assist radiologists and surgeons in determining disease margins and tissue homogeneity and support construction of brain atlases and pathology models.

***Index Terms*—** Texture descriptors, Vector quantization, Pattern analysis, Classification, Brain Images.

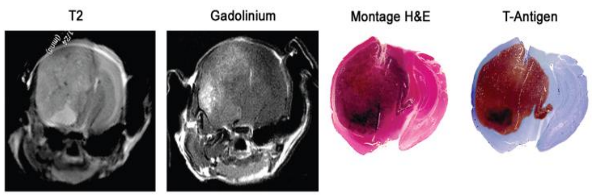
1. INTRODUCTION

It can often be difficult for clinicians to precisely assess the tissue composition of a lesion on the basis of radiographic appearance. Because this information is vital for determining accurate treatment and prognosis, such difficulty may necessitate more invasive examinations such as tissue biopsy, resulting in additional complications. Additionally, inability to determine tumor homogeneity may result in poor assessment of margins, possibly leading to incomplete resections and creating the potential for increased recurrence rates in patients with neoplasms of the brain or inadvertent excision of neighboring normal tissue.

Recent advances in medical imaging and computer-aided diagnostic modalities have improved the accuracy with which radiologists and surgeons can determine margins and texture composition. Nevertheless, manually assessing textural homogeneity and precise tissue margins remains rather subjective. In previous work, we have analyzed approaches for tumor segmentation using fuzzy-connectedness on entire images [1], keyblock-based classification of functional magnetic resonance activation patterns [2,3], and classification of brain images in Alzheimer's patients using a novel dynamic recursive partitioning approach [4]. In this paper, we focus on classification of individual regions of interest in a dataset of 23 post-T1 weighted gadolinium-enhanced Magnetic Resonance Imaging (MRI) slices and 83 pre-T2 weighted MRI slices of the brain of a single mouse, as shown in Figure 1. The mouse model used develops spontaneous CNS tumors due to the transgene, JC virus T-antigen, as described in [5]. Regions from MR images were extracted by domain experts and were supported by ground-truth histology data. In particular, we are interested in discriminating between 12 T1 and 28 T2 manually segmented ROIs from these images,

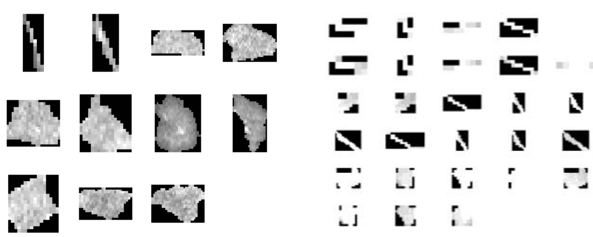
978-1-4244-2003-2/08/$25.00

representing various types of tissue: cerebrospinal fluid (CSF), gray matter, tissue necrosis, hippocampus tissue, and samples from three distinct regions of tumor with varying degrees of vascularization, neoplastic growth, and tissue necrosis. These ROIs are displayed in Figure 2. We wish to discriminate between individual tissue types as well as collectively classify tissue as normal (CSF, hippocampus, and gray matter classes) or abnormal (necrosis, tumor1, tumor2, tumor3). Among other applications, this research can be useful in determining disease properties and trajectories, such as tumor grade and stage, homogeneity of tissue, and composition of disease margins, all of which influence prognosis and treatment.



(a) **(b)** **(c)** **(d)**

Fig. 1: (a) Pre-T2 MR, (b) Post-T1 (gadolinium enhanced) MR, (c) hematoxylin and eosin histology, and (d) T-antigen histology images of a mouse with a large intracranial neoplasm. The dark area near the bottom of the image is necrosis.



(a) **(b)**

Fig. 2: (a) Post-T1 and (b) Pre-T2 ROIs prior to combination and nomalization. In order from top left, (a) 2 CSF ROIs, 3 Tumor1, 3 Tumor2, 3 Tumor3 and (b) 10 CSF, 2 Graymatter, 8 Hippocampus, 8 Necrosis.

2. BACKGROUND

The VQ approach is based on applying the keyblock image encoding [6,7] on image data to obtain compressed images. The keyblock approach decomposes each image into equi-size blocks and uses VQ to represent each block with the closest codeword from a codebook. First, given a fixed block size, each image is decomposed into a number of small blocks. Each small block contains features of the sub-area of its corresponding image. Based on such small blocks from different images, a codebook containing keyblocks is generated. In order to generate the codebook, the Generalized Lloyd Algorithm (GLA), which produces a "local optimal" codebook based on the nearest neighbor and the centroid conditions, is used. The algorithm is as follows:

Given a codebook $C_m = \{y_i\}$, an improved codebook C_{m+1} is generated by partitioning a training sequence T into cells R_i according to the Nearest Neighbor Condition:

$$R_i=\{x : d(x,y_i) \le d(x,y_j);\ \ \forall j \neq i\}$$

where $d(x,y)$ is the *distortion* between x and y and is computed via the Mean Squared Error. In other words, no two neighbors x and y may quantize to the same codeword if there exists a nearer neighbor of x that does not quantize to that codeword. The new codebook C_{m+1} is then set to the centroids of the new cells:

$$C_{m+1}=\{cent(R_i)\}$$

The algorithm then calculates the average distortion of C_{m+1}, denoted D_{m+1}, and stops if the fractional drop:

$$(D_m - D_{m+1}) / D_m$$

is below a user-defined threshold. Otherwise, the algorithm runs again.

To summarize, GLA starts with an initial codebook and converges upon a local optimum by iteratively applying the two conditions until the average distortion drops below a given threshold. Once an optimal codebook has been generated, each image is encoded using the codebook. In this manner, GLA is very similar to the process of k-means clustering. Once the codebook is computed, each image is decomposed into blocks, then for each block, the closest entry in the codebook is located and the corresponding index is stored. In such a way, each image is represented as a vector of frequencies of keyblock (codeword) appearance.

The Visual Vocabulary, or ViVo, approach [8] is an alternative procedure with several similarities to the keyblock approach. The general idea, as in the keyblock approach, is to decompose ROIs into a series of equally-sized blocks (or "tiles"). However, rather than using a codebook, the keyblock approach uses feature extraction techniques such as PCA or ICA to obtain ROI texture descriptors.

3. METHODOLOGY

Our classification approach is based on the same keyblock image encoding used in Vector Quantization [6] and described above. It is worth noting that we only use a single codebook for the entire dataset, rather than one codebook per class, in order to make our approach more general by modeling the conditions that such an analysis would be performed under on an entire image. Additionally, using different codebooks would introduce additional challenges: the codeword frequencies would no longer be useful features, since they would come from different codebooks. In this case, the codeword *contents* would need to be factored in with the frequencies. While we use only one codebook to analyze all ROIs, we generate a different one on each fold to ensure that the test data does not influence the trained codebook.

We apply the keyblock approach on each ROI to identify discriminative texture features within each region, optionally cropping or imputing the ROI with its mean to match the dimensions of other ROIs in the dataset. Normalization of ROIs may also be performed. Following data preprocessing, the keyword frequency vector is used as a representative feature vector of the image texture in classification. We do not perform the final step of vector quantization (substituting the codewords into the image to reconstruct a compressed image), as this is not necessary to capture texture information. We then employ the Histogram Model, which has been shown effective for texture classification in the literature [7], as a similarity measure in k-nearest neighbor classification, which determines the class of an ROI by a majority vote of its k nearest neighbors for a user-specified k. We additionally perform classification experiments directly on the feature vectors using the C4.5 decision tree learning algorithm, a support vector machine with a polynomial kernel (of degree 2), and a neural network with two hidden layers and a sigmoid transfer function for comparison. To accommodate the requirement of binary classification for a support vector machine, we split the original multi-class classification problem into binary one-vs-rest classification problems to ascertain SVM accuracy.

4. RESULTS

We performed leave-one-out classification experiments on a combined dataset of 12 post-T1 weighted (gadolinium enhanced) ROIs and 28 pre-T2 weighted ROIs extracted from 21 post-T1 and 83 pre-T2 slices of the brain of a single mouse afflicted with a large intracranial neoplasm. Images were registered prior to segmentation and normalized following combination. Segmentation itself was performed by domain experts and supported by histology data. A codebook was generated from the training data for each fold in the leave-one-out procedure; the test item was never included in codebook generation to ensure that the learned patterns were not influenced by the test item. The post-T1 ROIs were assigned labels “CSF”, “Tumor1”, “Tumor2”, and “Tumor3”, representing areas of cerebrospinal fluid, homogenous “typical” tumor tissue, heavily vascularized tumor tissue, and tumor tissue near an area of necrosis and edema. The tumor regions were clearly demarcated in the post-T1 images because gadolinium is highly sensitive to disruption of the blood-brain barrier, such as that which typically occurs in tumors. To take advantage of the imaging properties of T2 relaxation, we selected ROIs from the T2 image dataset in the following classes: “CSF”, “Graymatter”, “Hippocampus”, and “Necrosis”. These labels corresponded to areas of cerebrospinal fluid, normal gray matter tissue, a region of normal tissue located in the hippocampus, and a region of liquefactive necrosis near the lower central region of the tumor, respectively.

Results varied slightly over runs due to the locally optimal nature of GLA and the use of different starting conditions for codebook generation. Average kNN accuracies on the combined and individual datasets for values of k ranging from 1 to 6 are shown in the following table. In general, average results on the T1 dataset were lower than those of the combined dataset and fell off swiftly with increases in k, suggesting issues due to the small sample size, while average results on the T2 dataset were higher, suggesting that the T2 images may exhibit more consistent textural features, as they were taken primarily of normal tissue. The disparity between T1 and T2 results may also be due partially to the relative sizes of each dataset. It is also likely that normal regions of the brain (as were imaged in all classes of the T2 dataset except necrosis) exhibit greater textural homogeneity than regions of neoplastic tissue (as were imaged in all T1 classes except CSF), resulting in higher texture-based classification accuracy. Another possibility is that the relative dimensions of each T2 ROI influenced classification; however, neither cropping nor imputing missing values to obtain a uniform image size produced significantly different accuracies.

k	T1 + T2	T1	T2
1	**89.19%**	**81.82%**	**96.15%**
2	86.49%	63.64%	**96.15%**
3	62.16%	45.45%	84.62%
4	62.16%	36.36%	84.62%
5	72.97%	36.36%	80.77%
6	72.97%	18.18%	88.46%

Table 1: Average kNN classification accuracies for the T1, T2, and combined datasets.

In addition to computing class accuracies, we grouped classes into normal (CSF, Hippocampus, Gray Matter) and abnormal (Tumor1, Tumor2, Tumor3, Necrosis) categories and computed an ROC curve with an area of .8235, which is shown in Figure 3. We also classified ROIs using SVMs (polynomial kernel of degree 2, c=3), neural networks (sigmoid transfer function, 2 hidden layers, (out + in) / 2 hidden nodes), and C4.5 decision trees. These results are shown in Table 2.

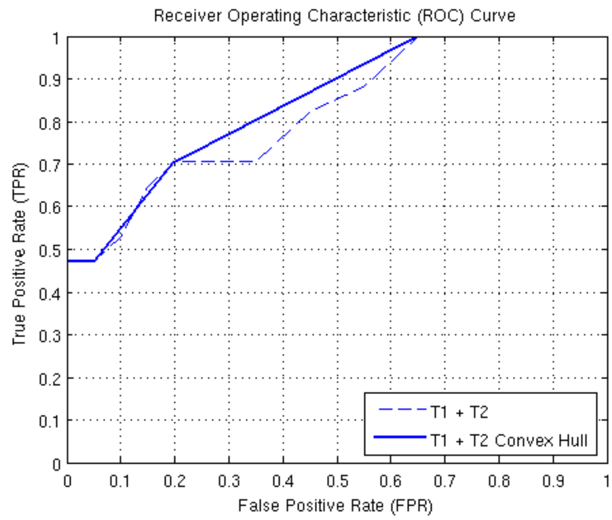


Fig. 3: An ROC curve plotting true positive rate against false positive rate on the combined T1 + T2 dataset.

Modality	*k*NN	SVM	Neural Net	C4.5
T1 + T2	89.19%	67.57%	86.49%	59.46%
T1	81.82%	72.73%	63.64%	9.09%
T2	96.15%	92.31%	96.15%	88.46%

Table 2: Best results of various classifiers.

Somewhat surprisingly, kNN achieved the best performance of all classifiers tested despite its relative simplicity. Also interestingly, SVMs performed fairly well on the individual datasets but produced poorer results when the datasets were combined, suggesting that two distinct hyperplanes must be learned for the two datasets. This may be possible to accomplish using a higher-dimensional kernel. As expected, a decision-tree approach performed poorly in this application, as the texture patterns distinguishing tissue type are unlikely to be separable by a rigid decision boundary.

5. CONCLUSION

Our methodology has the potential to assist in detection of tissue margins and determination of tissue composition through textural homogeneity, thus improving our understanding of brain structure and pathology and increasing the accuracy of diagnosis and staging. We demonstrate our methodology on 3 types of normal tissue and 4 types of abnormal tissue segmented from a mouse with a large intracranial neoplasm, obtaining average accuracies up to 89% on a combined T1 and T2-weighted dataset.

Opportunities for future work include utilizing two-step wavelet analysis in our approach, automatically segmenting ROIs using fuzzy segmentation techniques, integrating our methodology with brain image databases that we have previously developed, performing the same experiments on slices extracted from a larger number of mice or humans, including those with diverse tumor types, and applying these techniques more broadly to other medical imaging domains.

6. ACKNOWLEDGMENTS

This work was supported in part by NIH Research Grant #1 R01 MH68066-04 funded by NIMH, NINDS and NIA, and by NSF Research Grant IIS-0237921. The funding agencies specifically disclaim responsibility for any analyses, interpretations and conclusions.

7. REFERENCES

[1] Kontos, D., Wang, Q., Megalooikonomou, V., Maurer, A. H., Knight, L. C., Kantor, S., Fisher, R. S., Simonian, H. P., Parkman, H. P., A tool for handling uncertainty in segmenting regions of interest in medical images, International Journal of Intelligent Systems Technologies, Special Issue on "Intelligent Image and Video Processing and Applications: The Role of Uncertainty", Vol. 1, Nos. 3/4, pp. 194-210, 2006.

[2] Zhang, J., Megalooikonomou, V., An Effective and Efficient Technique for Searching for Similar Brain Activation Patterns, Proceedings of the IEEE International Symposium on Biomedical Imaging, Washington DC, Apr. 2007.

[3] Wang, Q., Megalooikonomou, V., Kontos, D., A Medical Image Retrieval Framework, Proceedings of the 2005 IEEE International Workshop on Machine Learning for Signal Processing (MLSP05), Mystic, Connecticut, Sep. 2005.

[4] Megalooikonomou, V., Kontos, D., Pokrajac, D., Lazarevic, A., Obradovic, Z., Boyko, O., Saykin, A., Ford, J., Makedon, F., Classification and Mining of Brain Image Data Using Adaptive Recursive Partitioning Methods: Application to Alzheimer Disease and Brain Activation Patterns, Presented at the Human Brain Mapping Conference (OHBM'03), New York, NY, Jun. 2003.

[5] Krynska, B., Otte, J., Franks, R., Khalili, K., and Croul, S., Human ubiquitous JCV(CY) T-antigen gene induces brain tumors in experimental animals, Oncogene, Vol. 18 no. 1, pp. 39-46.

[6] Mitra, S., Yang, S.Y., High Fidelity Adaptive Vector Quantization at Very Low Bit Rates for Progressive Transmission of Radiographic Images, Journal of Electronic Imaging, Vol. 8, pp.23-35, 1999.

[7] Zhu L., Rao, A., Zhang, A., Advanced Feature Extraction for Keyblock-based Image Retrieval, Information Systems, Vol. 27 no. 8, pp. 537-557.

[8] Bhattacharya, A., Ljosa, V., Pan, J., Verardo, M.R., Yang, H., Faloutsos, C., Singh, A.K., ViVo: Visual Vocabulary Construction for Mining Biomedical Images, Proceedings of the Fifth IEEE International Conference on Data Mining, Houston, TX, Nov. 2005.